6 Supplementary

Proof of Lemma 1:

Proof. In the 2-step procedure of Algorithm 1, S_0 is obvious the optimal solution to the sub-problem with parameter $|S_0|$, that is, $S_0 = S(|S_0|)$. Then for the second model selection step, $S_0 = S^*$ due to global optimality of S_0. □

To prove Thm.3, we need the following Finsler's lemma.

Lemma 8. *(Finsler) Let $x \in \mathbb{R}^n$, $B \in \mathbb{R}^{m\times n}$ and $Q \in \mathbb{R}^{n\times n}$ such that rank$(B) < n$, Q symmetric and positive semi-definite. Then the following two statements are equivalent:*

$$x'Qx > 0, \forall x \neq 0, Bx = 0 \iff \exists\gamma > 0 : Q + \gamma B'B \succ 0. \tag{19}$$

Proof of Thm.3:

Proof. Assume w.l.o.g. that $S = \{1, 2, ..., k\}$ consists of the first k nodes. Then $A \circ M$ exactly captures the adjacency matrix of the induced sub-graph:

$$A \circ M = \begin{pmatrix} A_S & 0 \\ 0 & 0 \end{pmatrix} \tag{20}$$

In the fashion, $diag\left((A \circ M)\mathbf{1}_n\right) - (A \circ M)$ captures the Laplacian matrix of S:

$$diag\left((A \circ M)\mathbf{1}_n\right) - (A \circ M) = \begin{pmatrix} L_S & 0 \\ 0 & 0 \end{pmatrix}. \tag{21}$$

By Lemma 2 and Rayleigh-Ritz theorem, we want the following to hold on L_S:

$$x'L_S x > 0,\ \forall 0 \neq x \in \mathbb{R}^k, x'\mathbf{1}_k = 0. \tag{22}$$

By Lemma 8, the above condition can be converted into:

$$L_S + \gamma \mathbf{1}_k \mathbf{1}_k' \succeq \epsilon I_k, \tag{23}$$

where $\gamma k \geq \epsilon$. Now we place this LMI back to the large matrix and notice the fact that:

$$diag(M\mathbf{1}_n) = \begin{pmatrix} kI_k & 0 \\ 0 & 0 \end{pmatrix}, \tag{24}$$

the equivalent LMI for the large matrix is:

$$diag\left((A \circ M)\mathbf{1}_n\right) - (A \circ M) + \gamma M \succeq \frac{\epsilon}{k} diag(M\mathbf{1}_n), \tag{25}$$

where $\gamma k \geq \epsilon$ should be satisfied. Let $\epsilon = \gamma k$, and the proof is done. □

Proof of Corollary 4:

Proof. Let $\gamma = \lambda_2(\Lambda_k)/k$. Then every S satisfying $Q(M;\gamma) \succeq 0$ and $diag(M)'\mathbf{1}_n = k$ is connected by Thm.3 and of size k. So $S \in \Lambda_k$.

On the other hand, for any $S \in \Lambda_k$, $\lambda_2(S) \geq \lambda_2(\Lambda_k) \geq \gamma k$. From the proof of Thm.3, the indicator matrix M corresponding to S satisfies $Q(M;\gamma) \succeq 0$ and $diag(M)'\mathbf{1}_n = k$. Proof is done.

□

Proof of Lemma 5:

Proof. If $M \in \{0,1\}^{n\times n}$, by constraints of Eq.(10), $M_{ij} = 1$ if and only if $M_{ii} = 1$ and $M_{jj} = 1$. Thus $M = diag(M)diag(M)'$ is rank-1.

On the other hand, if M is rank-1, or $M = ff'$, . Consider any two non-zero entries of f: $f_i = a \neq 0$, $f_j = b \neq 0$. Then by $M_{ij} \leq \min\{M_{ii}, M_{jj}\}$, we have $a = b$. So every non-zero entry of f is equal. The node with $M_{ii} = 1$ ensures that all non-zero entries of f is 1. Proof is done. □

Proof of Theorem 6:

Proof. For part (a), assume on the contrary that the support of $diag(M)$ is disconnected: $S = C \cup \bar{C}$, where $\bar{C} = S - C$. Let $|S| = k, |C| = k_1, \bar{C} = k_2$. W.l.o.g. assume $M_{11} = 1$, and C consists of nodes $\{1, 2, ..., k_1\}$.

Consider the $k \times k$ sub-matrix Q_S of Q corresponding to S, since the rest part are all 0. Now we use the vector $g = [\mathbf{1}_{k_1}; -\mathbf{1}_{k_2}]$ to hit Q_S:

$$g'Q_S g = g'\left(diag\left((A_S \circ M_S)\mathbf{1}_n\right) - (A_S \circ M_S)\right) g - \gamma g'\left(diag\left(M_S \mathbf{1}_n\right) - M_S\right) g \geq 0. \tag{26}$$

Note that A_S has the form:

$$A_S = \begin{pmatrix} A_C & 0 \\ 0 & A_{\bar{C}} \end{pmatrix}, \tag{27}$$

where the off-diagonal block is zero because by assumption C and $\bar{C}$ is disconnected. Then:

$$diag\left((A_S \circ M_S)\mathbf{1}_n\right) - (A_S \circ M_S) = \begin{pmatrix} \tilde{L}_C & 0 \\ 0 & \tilde{L}_{\bar{C}} \end{pmatrix}, \tag{28}$$

where $\tilde{L}_C$ is the Laplacian matrix of C weighted by M_C. Notice it still holds that $\tilde{L}_C \mathbf{1}_{k_1} = 0$. This means $g'\left(diag\left((A_S \circ M_S)\mathbf{1}_n\right) - (A_S \circ M_S)\right) g = 0$.

On the other hand, let $diag\left(M_S \mathbf{1}_n\right) - M_S$ be:

$$diag\left(M_S \mathbf{1}_n\right) - M_S = \begin{pmatrix} L_1 & L_3 \\ L_3' & L_2 \end{pmatrix}. \tag{29}$$

Using $g_1 = [\mathbf{1}_{k_1}; 0]$ and $g_2 = [0; \mathbf{1}_{k_2}]$ to hit Q_S will yield: $\mathbf{1}_{k_1}' L_1 \mathbf{1}_{k_1} = 0$ and $\mathbf{1}_{k_2}' L_2 \mathbf{1}_{k_2} = 0$. Apparently $g'\left(diag\left(M_S \mathbf{1}_n\right) - M_S\right) g \geq 0$ due to positive semi-definiteness of Laplacian matrix. If it's strictly positive, proof is done. Otherwise this means $\mathbf{1}_{k_1}' L_3 \mathbf{1}_{k_2} = 0$. Note that all entries of L_3 are either 0 or negative due to non-negativity of M_S. This means $L_3 = 0$, or equivalently $M_{ij} = 0$ for any $i \in C, j \in \bar{C}$. But this can not happen, because $M_{11} = 1$ and $M_{1j} \geq 1 + M_{jj} - 1 = M_{jj} > 0$ for any $j \in \bar{C}$. Contradiction!

Part (b) is straightforward by using $g = \mathbf{1}_C - \mathbf{1}_{\bar{C}}$ to hit Q_S. Proof is done. □

Proof of Proposition 7:

Proof. The proof is similar to the proof of Thm.6, by using $g = \mathbf{1}_{C_1} - \mathbf{1}_{C_2}$ to hit Q. □